Auburn Research: Student Symposium
Tuesday, April 9th, Auburn University Student Center

Please use this LaTeX template for your abstract. The Auburn University Library provides a wealth of resources on how to edit and format LaTeX documents at: https://libguides.auburn.edu/LaTeX.

INSTRUCTIONS

1. Edit this template in a LaTeX editor, like Overleaf. You should **only** change fields that have `% CHANGE THIS` above them. Do **NOT** delete any of the field names, e.g., `\textbf{Title:}`. Descriptions **cannot** be more than **2,000 characters**, *including spaces* (this will be automatically checked by a computer!). The title, author information, and affiliation are not part of the character count. Do not include figures or references in your abstract.

2. Proofread your abstract—it will appear as submitted!

3. Save a copy of this abstract template to your computer and label the file as `YOURLASTNAME.tex`.

4. Upload the file with your Student Symposium 2019 registration (instructions on the registration form).

5. Abstracts are due February 8, 2019, by 11:59 PM CST.

Title: Strike (with) a pose: neural networks are easily fooled by strange poses of familiar objects
Primary Author (and presenter): Alcorn, Michael, A.
Additional Authors: Li, Qi; Gong, Zhitao; Wang, Chengfei; Mai, Long; Ku, Wei-Shinn; Nguyen, Anh;
Department: Department of Computer Science and Software Engineering
College/School: Samuel Ginn College of Engineering

Description: Deep neural networks (DNNs) are increasingly common components of computer vision systems. When handling "familiar" data, DNNs are capable of superhuman performance; however, inputs that are dissimilar to previously encountered examples (but that are still easily recognized by humans) can cause DNNs to make catastrophic mistakes. Here, we present a framework for discovering DNN failures that harnesses 3D computer graphics. Using our framework and a self-assembled dataset of 3D objects, we investigate the vulnerability of DNNs to "strange" poses of well-known objects. For objects that are readily recognized by DNNs in their canonical poses, DNNs incorrectly classify 97% of their pose space. Further, DNNs are highly sensitive to slight pose perturbations; for example, rotating a correctly classified object as little as 8° can often cause a DNN to misclassify. Lastly, 75% to 99% of adversarial poses transfer to DNNs with different architectures and/or trained with different datasets.